Research Article
doi: 10.12973/eu-jer.10.1.285

European Journal of Educational Research
Volume 10, Issue 1, 285 - 298.
ISSN: 2165-8714
https://www.eu-jer.com/

A Critical Thinking Assessment Model Integrated with Science Process Skills on Chemistry for Senior High School

Kriswantoro*
Yogyakarta State University, INDONESIA

Badrun Kartowagiran
Yogyakarta State University, INDONESIA

Eli Rohaeti
Yogyakarta State University, INDONESIA

Received: September 12, 2020 ▪ Revised: December 8, 2020 ▪ Accepted: January 5, 2021

Abstract: Every school should be able to equip students to have the ability to integrate the knowledge gained with real life in responding to global challenges. Assessment of learning outcomes in the form of cognitive and skill aspects must go hand in hand. This study aims to produce: (1) a critical thinking model integrated with the science process, (2) construct validity and reliability instruments that measure the integrated critical thinking skills of science process skills in high school chemistry learning. This assessment model uses the Design and Development approach which refers to the Ellis & Levy model, namely: (1) problem (2) goal setting, (3) model design and development, (4) model testing, (5) evaluation of the test result model, and (6) application model. The validity assessment consisted of 3 measurement and measurement experts, 2 chemistry education experts, and 2 treatments. The trial sample described 289 high school students in Sleman Regency. Proving the validity of the content using the Aiken formula gets a value of 0.923 in the good category. Internal instrument reliability is included in the reliable category with a value of 0.886 (> 0.7), while the Composite Reliability ranges from 0.88 to 0.90. Confirmatory factor analysis showed that the value of the Root Mean Square Error of Approximation (RMSEA) was 0.021 <0.08, the Chi-Square obtained from <2df testing was 38.29 <2 (34) and the Goodness of fit Index (GFI) was 0.97. > 0.90 or a model declared according to the data obtained in the field and can be used in extensive measurements.

Keywords: *Assessment model, critical thinking, science process skills.*

To cite this article: Kriswantoro, Kartowagiran, B., & Rohaeti, E. (2021). A critical thinking assessment model integrated with science process skills on chemistry for senior high school. *European Journal of Educational Research, 10*(1), 285-298. https://doi.org/10.12973/eu-jer.10.1.285

Introduction

Learning is a process of reciprocal communication between teachers and students where teaching is carried out by teachers as educators and learning is carried out by students as learners. Chemistry learning is a process of interaction between students and their environment in order to achieve learning objectives. A learning process includes stages of learning, one of which is the assessment stage. Assessment is an evaluation process to determine the quality or success in the learning process in accordance with the desired learning objectives.

The purpose of learning chemistry is not only to produce individuals who master chemical concepts, but also have the skills to find these concepts and apply them in everyday life. In line with this, Marzano (2007) states that the goal of the learning process is clear and procedural knowledge. Clear knowledge is concise knowledge about the content of a science and can be developed through review and revision processes such as error analysis and identification of similarities and differences. Procedural knowledge is knowledge that is oriented towards skills, strategies and processes. The development of procedural knowledge is different from the development of declarative knowledge (Ismail & Jusoh, 2001). Procedural knowledge is developed by students through practice. This statement is consistent with its chemical properties as a product and as a process. Therefore, chemistry learning cannot be separated from declarative knowledge (chemistry as a product) and procedural knowledge (chemistry as a process).

Chemistry subjects are used as part of the secondary education curriculum, showing that chemistry has educational value in addition to its application to touch various aspects of human life. Studying chemistry is studying abstract chemical concepts. This view underlies the thought of curriculum development and chemical extension workers to design subject matter with broad theoretical and academic knowledge because it must cover all basic knowledge of chemistry. Chemistry was initially obtained and developed based on experiment (inductive) but in later developments

* **Corresponding author:**
Kriswantoro*, Yogyakarta State University, Educational Research and Evaluation, Yogyakarta, Indonesia.. ✉ kriswantoro.2017@student.uny.ac.id

chemistry was also obtained and developed based on the theory (deductive). From this, every school needs an integrated critical thinking assessment model of science process skills in learning and assessment in the classroom.

Critical thinking is part of higher-order thinking that is a major concern of education in the 21st century. According to some experts, one of Resnick's (1987) definitions of higher-order thinking skills is a complex thought process in describing the material, making conclusions, building representations, analyzing, and building relationships that involve the most basic mental activities. These skills are also used to underline high-order processes according to Bloom's taxonomy ladder. One of the most important skills that can motivate students in the 21st century is science process skills (Areesophonpichet, 2013). Chemistry lab activities are very interesting, so they can motivate students to work together in teams, look for ideas, and solve problems. In addition, practicum activities can form various skills in scientific work, including the skills to observe, ask, predict, formulate hypotheses, utilize tools and materials, process the data obtained, and describe the results obtained.

Every school must be able to equip students to have the ability to integrate the knowledge gained with real life in responding to global challenges. This is not an easy matter to make it happen because it takes skills to integrate it. South African Qualifications Authority (SAQA, 2014) states that integrated assessment is a form of assessment that involves all different types of assessment, such as written assessment (theory) and demonstration (practice) together with the aim of knowing the competencies of students.

Assessment is not a complement to the learning process, but is an integrated part of a learning process which becomes the basis for further improvement of the learning process. This is in line with Harrell (2010) that in an integrated curriculum, teachers must explicitly assimilate the concept of more than one discipline in the learning process. Furthermore, Lukum (2015) emphasizes that learning evaluation is an activity to evaluate things that are done in the learning process, including planning, implementation and the assessment process and its impact on students. Assessment is related to learning objectives and the learning process. The learning process will later be used as a reference by the teacher in carrying out the assessment. Then the results of the assessment will be used as a guide for evaluating the next learning process. Therefore, the learning process and the assessment of learning outcomes are two components that cannot be separated.

The achievement of student competencies after the learning process is the objective of the assessment in chemistry learning. Science process skills and critical thinking are competencies that must be achieved in learning chemistry. Science process skills are an inseparable part because they have an important role in mastering concepts in learning activities (Temiz et al., 2006). This is the reason why science process skills are so important to assess and develop. Durmaz and Mutlu (2014) explain that in the implementation of learning that places more emphasis on science process skills, learning outcomes will have a positive impact on student understanding. Shirban (2018) says that it should be noted that learning can be a fun process if students realize that learning is capable of producing something. Therefore, it is important to develop this assessment model in chemistry learning. Models are often seen as mediators that link theory and phenomena (Koponen, 2007; Rotbain et al., 2006). In science learning, Coll and Lajium (2011) report three main objectives of modeling, namely: (1) producing objects or concepts that are easy to understand; (2) provide stimulation to learning so as to increase the visualization of several events; and (3) provides an explanation of the science. The modeling process combines several characteristics of the fact model to the desired model (Gilbert & Justi, 2016). This condition requires them to be more active in continuous learning and also motivate teachers to further improve the quality of learning (Badu, 2012). Therefore, the aim of this research is to produce a critical thinking model that is integrated with science process skills and to prove the construct validity and reliability of the instruments developed.

Literature Review

John Dewey argues that critical thinking is essentially an active process, in which a person thinks things deeply, asks questions, finds the correct information rather than waiting for unclear information (Fisher, 2011). Critical thinking is a process where all knowledge and skills are directed to solve problems that arise, make decisions, analyze, and conduct research based on the data and information that has been obtained to produce the desired conclusions. Creative thinking is an ability that most of us who are not born natural creative thinkers. Special techniques are needed to help use our brains in different ways. Creative thinking can be in the form of imaginative thinking, generates many possible solutions, is different, and is lateral (Kusiak & Brown, 2007). Critical and creative thinking skills play an important role in preparing students to become good problem solvers and to be able to make mature decisions and conclusions and be able to be accounted for academically.

The focus of this research is high-level thinking as critical thinking. Critical thinking skills with good reason have become a concern and outcome of education (Ching & Fook, 2013). Many educational experts learn about critical thinking skills, resulting in different understandings of critical thinking. Dewey (1993) calls critical thinking reflective thinking. Critical thinking according to Fisher (2011) is an active, persistent, and careful consideration of beliefs that are taken for granted, in terms of supporting reasons and conclusions.

Based on this understanding, high-order thinking skills as Critical can be interpreted as thinking skills which involves cognitive processes and invites students to think reflectively in problem solving. The instrument for assessing critical thinking skills developed in this study will adopt higher-order thinking assessment aspects as critical thinking according to Bloom as seen based on higher-order thinking indicators or HOTS which includes analyzing (C4), evaluating (C5), and creating (C6) (Ramadhan et al., 2019) . The indicator is modified according to the chemistry material for class XI SMA which is applied in the research.

Developing science process skills are one of the important goals of learning chemistry. Science process skills is part of the mental and physical skills scientists use to solve problems, understand content in depth, develop an attitude of responsibility, and improve the learning experience of students (Gurses et al., 2014; Özgelen, 2012). Science process skills are a key aspect in studying scientific knowledge because this knowledge is obtained from the results of investigations. In practice, science process skills development also aims to support critical thinking skills and develop problem solving. Al-Rabaani (2014) science process skills is able to increase knowledge to build student abilities and develop the knowledge gained. Akinbobola and Afolabi (2010) explain that cognitive abilities and problem-solving skills can be realized by science processing skills.

Science process skills need to be applied because it will indirectly involve students in inquiry activities and be able to apply integrated assessment of process skills in learning (Leonor, 2015). In general, the evaluation system and learning model with the application of the science process skills assessment proved to be quite effective in improving student learning outcomes and science process skills (Hidayati et al., 2013). According to Özgelen (2012), this condition occurs because science process skills and critical thinking have a relationship with the cognitive domain of students. Basic and integrated skills are part of science process skills (Akinbobola & Afolabi, 2010; Al-Rabaani, 2014; Chabalengula et al., 2012; Karamustafaoğlu, 2011; Karsli & Şahin, 2009; Keil et al., 2009; Leonor, 2015; Özgelen, 2012; & Subali, 2010). Therefore it is said that chemistry subject teachers must be proficient in applying science process skills at various levels and have broad knowledge and understanding.

Methodology

This study used a Design and Development (D&D) approach. This approach is a systematic study related to the design, development and evaluation process with the aim of building an empirical basis for creating new instructional products, tools and models (Klein, 2014; Richey & Klein, 2007; Spector et al., 2014). The application of the Design and Development approach in this study uses a type of Research Model. The product produced in the implementation of this research is an Integrated Critical Thinking Assessment Model for Science Process Skills for high school students in chemistry learning.

Sample and Data Collection

The sample used was the students of class XI IPA SMA who were taking an odd semester learning for the 2019/2020 school year. The taking of research subjects was seen from the National Exam scores for the 2018/2019 academic year seen from the average ranking of the highest, medium, and low grades of chemistry subjects in Sleman Regency with a total of 289 students. The data collection technique in this study was a test. The instrument for the data collection test in this study was a set of questions. The form of the critical thinking assessment test instrument is integrated with the ability of chemical science processes in the form of essay questions that have been arranged based on the grid and the material to be tested. All test subjects were asked to take the test according to the test equipment received. The following is an example of a question form that has been developed.

The following diagram shows the pH value data for four solutions (A, B, C, and D) that have the same concentration.

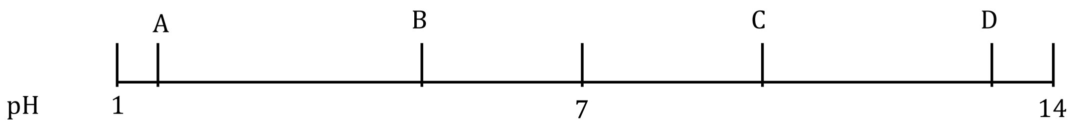

Figure 1. The pH value of the solution

The four solutions are phosphoric acid, ammonia, cyanide acid, and calcium hydroxide. (1) Write down the ionization reactions for all these solutions, and (2) Based on the diagram, make a table to classify these solutions into acids (monoprotic, diprotic, or tripotic) or bases (monohydroxy or polyhydroxy) and their valence acids or bases!

Analyzing of Data

In the design phase, a literature study and synthesis of findings were carried out that supported the development of an integrated critical thinking ability assessment model integrated science process skills of high school students in chemistry learning. Arranging a grid of question items is necessary in order to make it easier to make good item questions. This research begins with compiling a grid of questions based on the Core Competencies and Basic Competencies contained in the applicable learning syllabus. The lattice item is a reference in question writing, so that by referring to the lattice, the compiler will produce questions with relatively the same difficulty level. In addition, in the initial model phase a prototype 1 was produced and then continued in the model testing phase.

To prove the validity of this study, it was analyzed by looking at the readability of the content validity using Aiken's formula and construct validity. Proving the validity of the content in this study was analyzed from the score given by expert judgment involving 3 experts from the field of measurement (test construction experts), 2 material experts (chemistry), and 2 practitioners. The integrated critical thinking assessment skills assessment model that has been compiled in this study is then validated so that the assessment model truly describes what is being measured. The test items developed are based on those that are proportionally distributed according to the material and theory listed in the curriculum so that the theoretical validity meets the requirements. Proving the validity of the content in this study was analyzed from the scores given by experts from the field of measurement (test construction experts) and material experts (chemists). The validity coefficient of around 0.7 is still acceptable and considered satisfactory (Aiken, 1980). The results of the judgment from the judgment are then analyzed using the Aiken formula (Azwar, 2015) in equation 1.

$$V = \frac{\sum s}{[n\,(c-1)]} \quad (1)$$

Information:

s = r - lo

lo = the lowest number of validity assessments

c = the highest validity score

r = the score given by the assessor

Construct validity is a used for set of measuring instruments that have multiple indicators. Valid measuring instruments obtain consistent data even though the indicators used vary. In order to avoid overlapping measurements between indicators, there needs to be a conceptual definition with clear boundaries. To test the instrument construct, Confirmatory Factor Analysis (CFA) from Kirsch and Guthrie (1980), Shavelson and Stanton (1975) was carried out. Reliability is the extent of the reliability of the test in measuring the desired latent constructs. The measurement model reliability assessment includes Internal Reliability (Cronbach's alpha), Composite Reliability, and Average Variance Extracted. Internal reliability can be obtained with the help of SPSS, while CR and AVE are found by equations 2 and 3.

$$AVE = \frac{\sum K^2}{n} \quad (2)$$

$$CR = \frac{\sum K^2}{[(\sum K)^2 + \sum(1-K^2)]} \quad (3)$$

Information:

K = *factor loading*

n = the number of items in the model

Findings / Results

Results of the Design Stage

This phase is the phase of designing and developing a model or prototype. At this stage, a literature study and synthesis of literature study findings are carried out to support the development model. Furthermore, this phase of the initial design model prototype 1 has been completed and then continued in the model testing phase. The literature study carried out in this phase is to examine the concepts of critical thinking as higher order thinking (HOTS), Science process skills (SPS), characteristics, and assessment methods as a basis for developing an initial design model or prototype 1.

After examining the concept of critical thinking integrated scientific process skills, research insights are obtained related to the theory, procedures, steps, and methods used in development research. Through this activity, it is hoped that an empirical study on the implementation of integrated critical thinking assessment skills in chemistry lessons in high school is expected to be obtained. Based on the study of relevant research results, what has been explained shows that the integrated critical thinking assessment of science process skills is very supportive of learning in schools. To find out the abilities of these students, it can be done by applying an integrated critical thinking assessment model of science process skills (SPS). This critical thinking is defined as thinking skills that involve children's cognitive and reflective thinking about a problem. The instrument for assessing critical thinking skills developed in this study will adopt higher-order thinking assessment aspects as critical thinking according to Bloom as seen based on higher-order thinking indicators or HOTS which includes analyzing (C4), evaluating (C5), and creating (C6) . The purpose of science process skills is to improve students' thinking skills and make learning more active. Science process skills in this study includes activities (1) observe, (2) communicate, (3) inference, (4) creates data, and (5) designing experiments.

Results of literature study and preliminary activities, then the problem and needs analysis was carried out. The problem analysis shows that the test instrument that has been developed by the teacher has not measured the critical thinking skills integrated with the science process skills of students. The assessment model for the integrated critical thinking test instrument of science process skills developed by the teacher still measures at a low stage. The assessment tools designed by the teacher have not been able to measure and improve thinking skills.

Problem analysis was carried out to identify a research model that was able to improve students' critical thinking integrated science process skills in chemistry learning. Analysis of needs in this study obtained results (1) it is necessary to develop teacher skills in making instrument models capable of measuring critical thinking skills integrated with science process skills, and (2) The assessment model developed is able to capture students' weaknesses in thinking.

Model Development Results

The result of this development is a product in the form of a model equipped with a critical thinking test kit integrated science process skills in chemistry subjects for class XI SMA even semester of the 2019/2020 school year. In the early draft model development stage, the activities carried out were developing a draft prototype model and model validation. Model development aims to produce a model that can be applied properly and effectively. The prototype model consists of scoring instruments and guidelines as well as how to interpret the results of the assessment. The planning of the prototype model is carried out through the activities of designing the model, including planning the model objectives, model characteristics, components, instruments, syntax, and implementation guidelines. The initial draft of the assessment model prototype is then reviewed by involving supervisors and experts in the field of chemistry and research and development.

Based on the review by the experts, several suggestions and input were obtained regarding the design of the assessment model prototype which were grouped as follows: (1) the objectives must be clearly arranged and in accordance with the assessment principles; (2) characteristics should be made specifically and clearly in order to have advantages; (3) model components must reflect the principles of assessment; (4) instruments need to be classified, observed and improved so that they are clearer and more systematic; and (6) model implementation guidelines need to be prepared in a practical, complete, and clear manner. After being repaired and revised, the prototype of the assessment model is then validated by experts in order to obtain a model that is suitable and suitable for use.

The next activity is model validation, model validation activities are carried out by measurement experts and chemistry education experts, namely 2 experts in the field of chemistry and 3 experts in the field of measurement who have experience in writing and development. The prototype model that has been validated by the expert is then analyzed by the researcher for legibility testing. Aspects that are considered in this analysis include aspects of material, construction, and the language used. The following is presented in Figure 2.

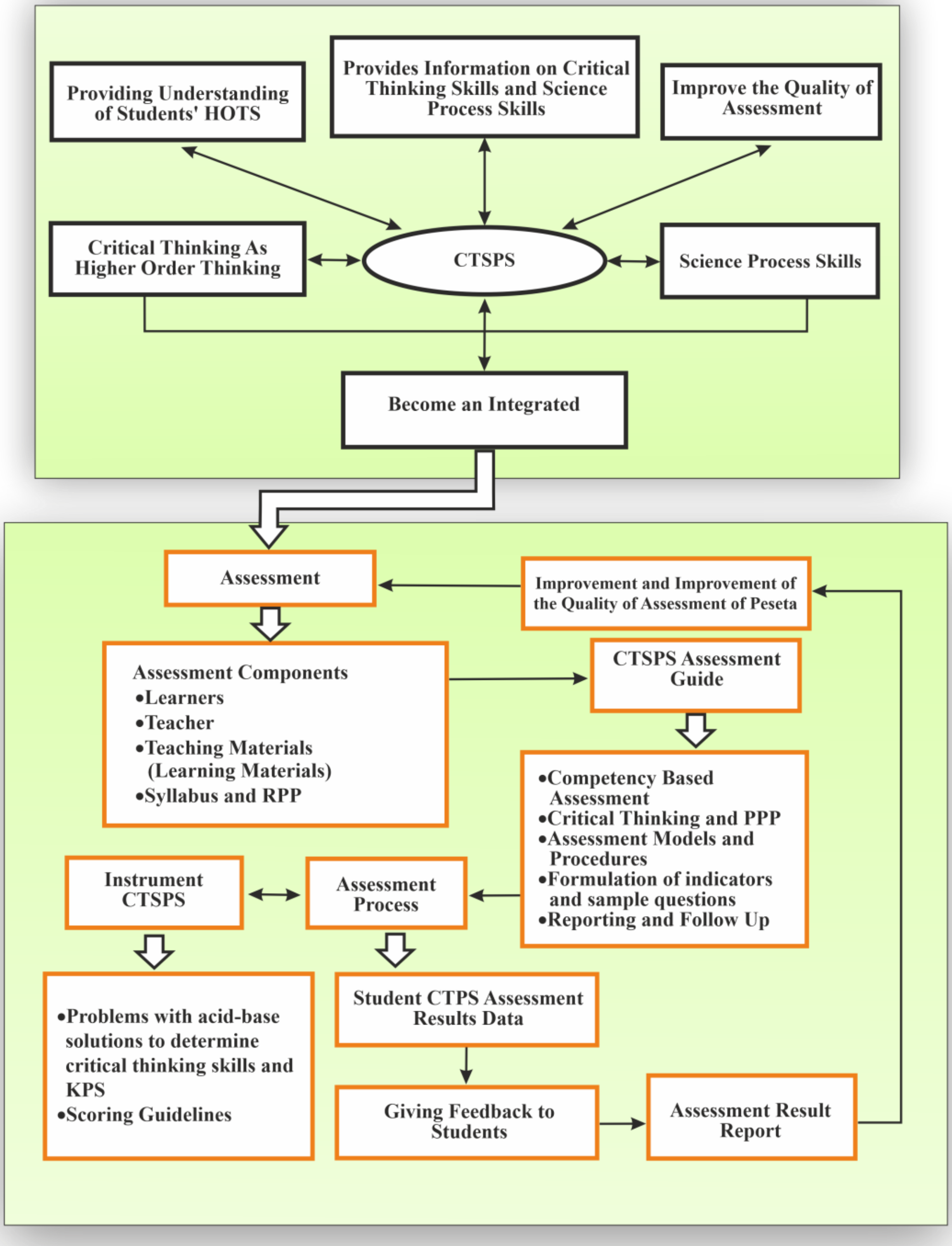

Figure 2. CTSPS scoring model

Activities in this validation aim to provide content validity. At this stage the experts assess whether the assessment model developed has met the criteria or not. In general, the review referred to includes a review of the test material, a review of test construction, and a study of language. Each expert provides an assessment for the accuracy of the grain in measuring the acid-based solution material. The score given has a range of 1-4 (irrelevant, less relevant, relevant, and highly relevant). The results of the validation from the experts were in the form of assessments and input both orally and in writing.

The decision to validate the model is carried out by comparing the Aiken index calculated for each assessment item with the reference value of the Aiken Table. Based on the Aiken table for statements validated by seven experts, the score criteria / branches used were 4, and the significance used was 0.05, then the Aiken reference value was 0.86. The results showed that the Aiken index was 0.923. Each statement for the assessment that has an Aiken index calculation result of at least 0.86 can be concluded as a valid item or meet the content validity.

Product Trial Results

Further analysis was carried out by Confirmatory Factor Analysis (CFA). CFA is to make a confirmed model formulation derived from theory. CFA has two stages, namely: (1) model analysis (2) model testing. Based on the instrument construct that has been developed, it can be seen that there are 2 factors. Based on this framework, the character construction is more suitable using the second order CFA can be seen in Table 1.

Table 1. Instrument Matrix

Aspects of Critical Thinking as Higher-Order Thinking		**Aspects of Science Process Skills**				
		Observe	**Communicate**	**Inference**	**Create data**	**Designing experiments**
Analyze (C4)	A1			10b (1.1)		
	A2	6a (1.5)	5a (1.6)			
	A3	10a (1.1)		6c (1.5)		1d (1.2)
Evaluate (C5)	B1		7 (1.3)	5b (1.6)		
	B2	2 (1.7)	6b (1.5)	4c (1.9)		4b (1.9)
Create (C6)	C1	8a and 8b (1.4)		1b and 1c (1.2)		
	C2		5c (1.6)		3b (1.2)	9b (1.8)
	C3	1a (1.2)	4d (1.9)	3a (1.2) and 9a (1.8)	5a (1.9)	

Description:

Theory :

1.1 Criticizing the ideas and workings of acid and base reactions according to the Arrhenius, Brønsted-Lowry, and Lewis acid-base theory
1.2 Describe the special characteristics of acid and base reactions
1.3 Sorting the types and properties of acid and base solutions
1.4 Analyzing acidity (pH)
1.5 Connect the components and the workings of the acid / base solution with the degree of ionization (α) and the acid constant (Ka) or the base constant (Kb)
1.6 Check reactions and estimate the pH of a solution using several indicators
1.7 Describe and identify the different types of reactions in solution
1.8 Determining the levels of various acid and alkaline solutions by titration method
1.9 Predict experimental results and conclude results based on experimental titration data

The constructs can be grouped into unidimensional and multidimensional constructs. From the previous test, namely the dimensional data test, it shows that the data are multidimensional. The construct of critical thinking variables and science process skills is hypothesized to consist of two factors, namely factor A (acidity level) and factor B (reaction equation). The results of the model fit test with the test data and the confirmatory factor analysis, path diagram using LISREL, the result of the path analysis construction model is obtained. It can be seen that the Estimated Output in Figure 3.

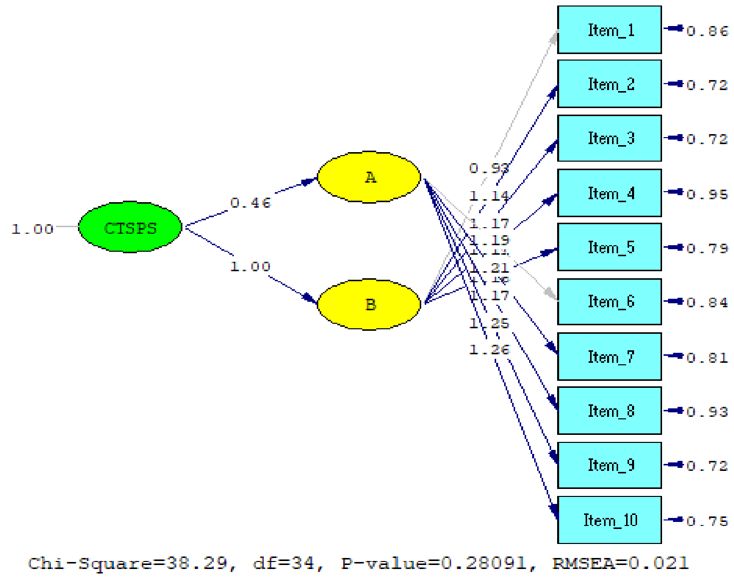

Figure 3. Structure of Output Dimensions Estimation Assessment Critical Thinking Integrated Science Process Skills

The most important part to note from Figure 3 is model fit. There are many criteria for model fit, more details can be seen in Table 2.

Table 2. Model Fit Test Results

Number	Indicator	Score Benchmark	Score Acquisition	Model fit
1	Chi-Square	< 2df	38.29 < 2(34)	Good
2	Probability (p-value)	≥ 0.05	0.28091	Good
3	Root Mean Square Error of Approximation (RMSEA)	≤ 0.08	0.021	Good
4	Root Mean Square Residual (RMSR)	≤ 0.08	0.06	Good
5	Normed Fit Index (NFI)	≥ 0.90	0.99	Good
6	Comparative Fit Index (CFI)	≥ 0.90	1.00	Good
7	Incremental Fit Index (IFI)	≥ 0.90	1.00	Good
8	Goodness of Fit Index (GFI)	≥ 0.90	0.97	Good
9	RFI	≥ 0.90	0.98	Good
10	Adjusted Goodness of Fit Index (AGFI)	≥ 0.90	0.96	Good
11	Parsimony Goodness of Fit Index (PGFI)	≥ 0.60	0.60	Good

Referring to the opinion of Joreskog and Sorbom (1989) a model is said to meet the criteria for the appropriate model if it meets at least three indexes of the fit model. The fit model index is the Chi-square test ($\chi 2$), RMSEA, P-Value and GFI. Table 2 shows that in general the requirements for Goodness of Fit have been met because the values obtained are within the required interval, so that it says that the model obtained is appropriate, that is, the construct of the instrument is good and it can be used to spread the model in this case the measurement.

Furthermore, to determine the significance of the influence between variables, it is seen from the output t-values in Figure 4.

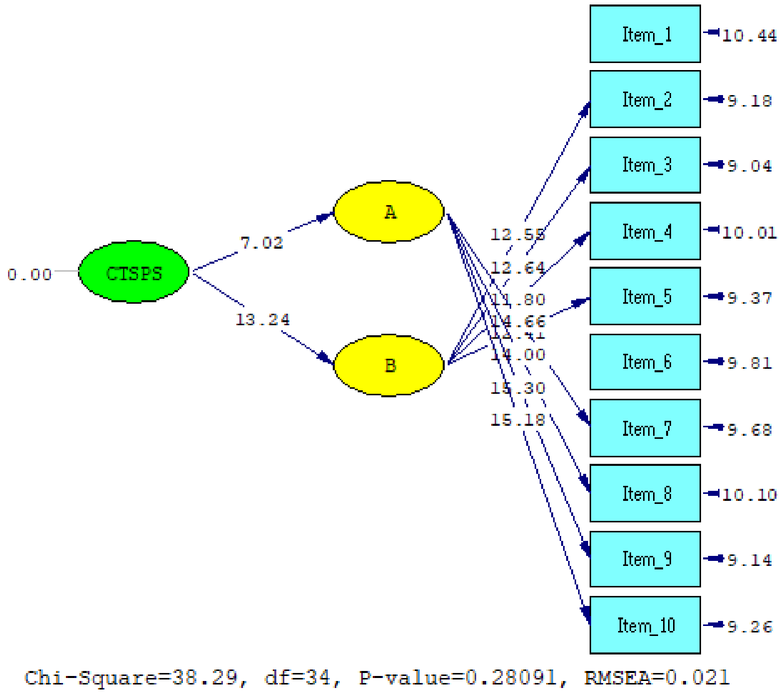

Figure 4. Output t-values

After getting a good model, the next step is to test the significance of the parameters using the t statistical test. If the t value of the output is red, it is not significant, whereas if the t value of the output is black, it is significant. Figure 4 shows that the effect of factor A (acidity level) and factor B (reaction equation) is significant, as indicated by the t value which is not colored in red. Path testing can be done by comparing the t value with the critical t. The critical t value has been set at 1.96 at alpha 0.05 (5% significance level). The analysis result shows the t-value> 1.96. Then the analysis is carried out by looking at the factor loading the effect of factor A (acidity level) and factor B (reaction equation) on each item. After analyzing using LISREL, the output is obtained in Figure 5.

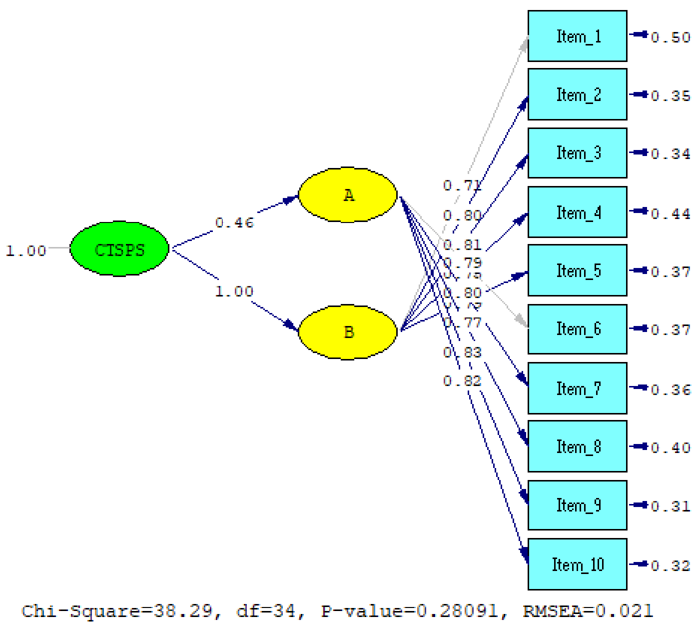

Figure 5. Output Path Diagram of Standard Solution Analysis Results

Once it is known that the model is fitted, further analysis is carried out to test the validity of the construct. The construct validity test is carried out by paying attention to the standard loading factor value of each indicator or dimension. If the standard loading factor value ≥ 0.40 then it is declared valid (Retnawati, 2016). The analysis shows the value of the standard loading factor ≥ 0.40 so it can be declared valid. From the standard solution output all indicators have a loading factor of more than 0.40; namely 0.46 for indicator A (acidity level) and, 1.00 for indicator factor B (reaction equation). Thus, it can be said that the two indicators, namely the level of acidity and the reaction equation are valid for describing the CTSPS model. Furthermore, based on the standard solution output above the factor loading value for all grains is more than 0.40. Convergent validity in this study is fulfilled because the entire Average Variance (AVE) is more than the minimum limit of 0.5 (Table 3). Meanwhile, construct validity is also fulfilled, this is because all the criteria for the model fit test / goodness of fit are met (Table 2). So it can be concluded that all items are valid for measuring indicators.

Reliability is the extent of the reliability of the measurement model in measuring the intended latent constructs. The assessment of the reliability of the measurement model includes Internal Reliability, Composite Reliability, and Average Variance Extracted can be seen in Table 3.

Table 3. Reliability analysis results

Latent Variable	Item Number	Loading Factors (min 0.4)	Composite Reliability (CR) (min 0.6)	Average Variance Extracted (AVE) (min 0.5)
A	Item 6	0.79		
(acidity level)	Item 7	0.80		
	Item 8	0.77	0.90	0.64
	Item 9	0.83		
	Item 10	0.82		
B	Item 1	0.71		
(reaction equation)	Item 2	0.80		
	Item 3	0.81	0.88	0.59
	Item 4	0.75		
	Item 5	0.79		
Internal Reliability (Cronbach's alpha) :		0.886		

Internal reliability of the instrument is included in the reliable category with a value of 0.886 (> 0.7) while reliability for each construct / Composite Reliability ranges from a value of 0.88 to 0.90 more than the minimum limit of 0.6. AVE is calculated as the root mean standardize loading factor divided by the number of indicators. AVE is able to show the ability of latent variable values to represent the original data score. The greater the AVE value, the higher the ability to explain the value on indicators measuring latent variables. The cutoff value of AVE that is often used is 0.50 where the AVE value is at least 0.50, the results of the analysis show that a good convergent validity is 0.59 and 0.64 > 0.50 which means that the probability of indicators in one construct entering other variables is lower (less 0.50) so that the probability of the indicator converges and enters a construct where the value in the block is greater than 50%. This indicates that this instrument has a degree of consistency that is high if used on other occasions even with different respondents.

Discussion

The final product in this study is a critical thinking assessment model integrated science process skills in chemistry subjects in high school grade XI. As stated by Cantos et al., (2015) that the assessment is expected to reflect the ability of students as a whole, both in terms of knowledge, attitudes, and skills (Crawford & Kirby, 2008; Kasilingam et al., 2014) and able to stimulate students to optimize their potential (Abell, 2006; Ketterlin- Geller, 2005). The integrated critical thinking assessment model of science process skills in high school chemistry subjects in grade XI has passed a series of tests consisting of qualitative analysis, validation, testing, and measurement at the time of model deployment. Amelia and Kriswantoro (2017) say that the test instrument that is compiled must meet the criteria as a good measuring tool in order to provide an overview of the abilities and competencies of students. The initial development consisted of 10 test items. The results of the test items showed that of the 10 test items had met the requirements in the analysis.

As previously explained, model testing is carried out, namely testing by experts and testing empirically through trials. The results of the expert's assessment of the preliminary design assessment model indicate that the developed model has an integrated critical thinking aspect and science process skills accordingly. An instrument is said to be valid if the expert believes that the instrument measures the things to be measured (Ramadhan et al., 2019). In addition, the 10 items contained in the development model will be valid and accepted according to input and suggestions from experts. The input and suggestions were corrected according to the notes by the experts. After being repaired, the model is feasible to prove with the Aiken index magnitude of 0.923 and can be used at the next stage, namely empirical testing through trials. These results are labeled in this study as the Revised I CTSPS Assessment Model. Field trials in 6 schools spread across several sub-districts in Sleman Regency. The analysis results show that of the 10 items developed has a loading factor value. The criteria used are RMSEA 0.021<0.08, Chi-Square obtained from testing <2df, namely 38.29<2

(34) and Goodness of fit Index (GFI) 0.97>0.90 (Kartowagiran et al., 2020; Schermelleh-Engel et al., 2003). In general, the requirements for Goodness of Fit have been fulfilled because the value obtained is within the required interval, so it is said that the model obtained is fit, thus it is concluded that the model used is fit, which means that the instrument construct is good and can be used in terms of model deployment this measurement.

Richey (2005) emphasizes that validation is seen from internal validation (confirmation of components and processes) and external validation (validation of the impact of using models). Validation that involves expert review, expert review is the process by which the expert judges in terms of components, structure, and future use. Basically, the review process and assessment input are based on predetermined criteria, and the subsequent model revisions are based on data input and suggestions from experts. This type of validation procedure can be referred to as a type of formative evaluation.

In line with Tracey and Richey (2007), the Delphi technique is used because it is more representative as a validation process that involves experts to criticize and assess the components and overall structure of the developed model. More specifically, it can be concluded that there are two aspects of the Delphi process which prove to be very valuable in this study. First, this technique proved successful because of the qualification of the assessment. The assessment panel has expertise not only in one area of expertise, but in model construction and use. Experts are an important part of the internal validation process. Expert reviewers were given one week to review each round, answer several open-ended questions in the first round and produce the most significant model revisions. This gives each expert the opportunity to rate and comment within a flexible time frame. Inputs and suggestions are very important to be reviewed in the next revision.

Further study is an empirical study that further describes the processes involved in construction or refinement. This research can explain the process and construction of the developed model. First, this approach makes it possible to finish in the creation of a new model with an appropriate level of clarity with a focus on several aspects. Second, the results of the model can be easily mastered by users, in this case the teachers in schools. This study produces a model that has proven the validity of its content and can be used by teachers who cannot be separated from the context, content, and students.

Temiz et al. (2006) emphasized that science process skills are an inseparable part and at the same time have a central role in developing students' conceptual understanding in learning activities. This condition makes the reason why science process skills are very important to be developed and assessed. In addition, Durmaz and Mutlu (2014) explained that if the learning that is carried out is emphasized more on the development of science process skills, it will have a positive impact on student learning outcomes. Therefore, the assessment of science process skills is an important component in learning chemistry. This condition is because the assessment can encourage students to be more active in learning continuously and also encourage teachers to further improve the quality of the learning process (Badu, 2012).

Improved thinking skills are not only limited to understanding concepts, but also involve thinking skills and dealing with real-life problems related to assessing information and arguments in a social context and making life decisions (Bailin, 1987). The relatively low score of critical thinking skills is because students often experience difficulties in choosing, formulating, and connecting experimental findings that can be used as evidence in supporting arguments (Katchevich et al., 2013).

There are various ways that can be done to promote critical thinking skills integrated with science process skills. Strategies commonly used by teachers in class to improve these abilities are through learning and assessment. Ideally, the assessment activities carried out should be in line with the learning approach used. Assessment of critical thinking skills integrated with science process skills will be appropriate if the learning approach used already involves these abilities. If the learning that is carried out does not encourage this ability, it does not mean that the integrated critical thinking assessment of science process skills cannot be carried out.

Runco (1990) says that the thinking ability of students is significantly influenced by the opportunities given to them. Teachers who are used to providing opportunities for students to do unfamiliar tasks and are able to hone their thinking skills certainly encourage students' critical thinking skills more than teachers who only give routine assignments. The important role of the teacher in facilitating the critical thinking skills of students must also coincide with the critical thinking ability of the teacher itself. Basically the teacher is a role model in creating critical students. This is in line with the opinions of Munawaroh et al. (2018) and Runco states that critical students can be formed if teachers who teach are also able to think critically. Not only improving learning patterns in all aspects, but also changing the mindset of teachers, improving learning patterns in educational staff producing institutions, and support of families and the surrounding environment and teachers must also be able and know the appropriate assessment model for students.

Conclusion

The conclusion that can be drawn from the results of developing an assessment model in this study is that the developed model has good characteristics and qualities as a model that is useful and meets the requirements used to measure the abilities of students. This is evidenced from the data analysis results which ensure that the CTSPS model has met the content validity of the expert judgment and the construct validity and reliability with confirmatory factor analysis of the field trial data. It cannot be denied that there are many shortcomings and weaknesses in this research that can be developed by the teacher or researcher who will come for the perfection of the future model, one of which is that the material being developed is still narrow. The CTSPS assessment model is still rarely applied in learning in the assessment process, so that students are not used to it and still find it difficult to do critical thinking tests integrated science process skills.

Recommendations

The assessment model developed is still limited to one basic competency, so it is very necessary for the development of teachers and other researchers to formulate other competencies that are suitable in chemistry learning. This is a very important point and should be carried out in a sustainable manner so that many items will be collected which can create a regional level question bank, especially in chemistry subjects. Other researchers who are interested in similar topics should develop an integrated critical thinking assessment of science process skills in other fields or materials because this is very useful for teachers and students in particular and the development of the world of education in general.

Limitations

There are several limitations that can be used as learning for further research improvement. Among other things (1) the models and instruments developed were still limited to students in one basic competency in class XI high school chemistry learning; and (2) the dimensional structure of the tested instrument construct using the CFA model assuming this instrument has been constructed properly based on the test specification table and through the item validation review by the expert judgment. This focus is to see the quality of the developed models and instruments by proving their validity and reliability.

Acknowledgements

Thank you to the Yogyakarta State University Postgraduate Program Education Research and Evaluation Study Program which has provided a wealth of research experience and enabled writers to learn more. Thank you to the chemistry teachers and high school students in Sleman Regency who have helped and contributed to the data collection of this research.

References

Abell, M. (2006). Individualizing learning using intelligent technology and universally designed curriculum. *The Journal of Technology, Learning and Assessment*, *5*(3), 4-20.

Aiken, L.W. (1980). *Content validity and reliability of single items or questionnaires*. Pepperdine University.

Akinbobola, A. O., & Afolabi, F. (2010). Analysis of science process skills in West African senior secondary school certificate physics practical examinations in Nigeria. *American-Eurasian Journal of Scientific Research*, *5*(4), 234-240.

Al-Rabaani, A. (2014). The acquisition of science process skills by Omani's pre-service social studies' teachers. *European Journal of Educational Studies*, *6*(1), 13-19.

Amelia, R. N., & Kriswantoro, K. (2017). Implementation of Item Response Theory for Analysis of Test Items Quality and Students' Ability in Chemistry. *Journal of Chemistry and Chemical Education/ Jurnal Kimia dan Pendidikan Kimia*, *2*(1), 1-12. https://doi.org/10.20961/jkpk.v2i1.8512

Areesophonpichet, S. (2013). *A development of analytical thinking skills of graduate students by using concept mapping*. The Asian Conference on Education.

Azwar, S. (2015). *Reliabilitas dan validitas* [Reliability and validity]. Pustaka Pelajar.

Badu, S. Q. (2012). Implementasi Evaluasi model Kirkpatrick pada perkuliahan masalah nilai awal dan syarat batas [Implementation of kirkpatrick's evaluation model in the course of initial value and boundary condition problems]. *Journal of Educational Research and Evaluation/ Jurnal Penelitian dan Evaluasi Pendidikan, 16,* 102-129. https://doi.org/10.21831/pep.v16i0.1108

Bailin, S. (1987). Critical and creative thinking. *Informal Logic, 9*(1), 23-30. https://doi.org/10.22329/il.v9i1.2656

Cantos, A. E., Alday, M. G., Alog, K. J., Asi, K. J., Calacal, R. H., & Britiller, M. C. (2015). Changing learning needs of student nurses: input to the nursing curriculum. *Asia Pacific Journal of Multidisciplinary Research, 3*(3), 108-119.

Chabalengula, V. M., Mumba, F., & Mbewe, S. (2012). How pre-service teachers' understand and perform science process skills. *Eurasia journal of mathematics, science and technology education, 8*(3), 167-176. https://doi.org/10.12973/eurasia.2012.832a

Ching, H. S., & Fook, F. S. (2013). Effects of multimedia-based graphic novel presentation on critical thinking among students of different learning approaches. *Turkish Online Journal of Educational Technology*, *12*(4), 56-66.

Coll, R. K., & Lajium, D. (2011) Modeling and the future of science learning. In M. S. Khine & I. M. Saleh, (Eds.), *Models and modeling: Cognitive tools for scientific enquiry*. Springer. https://doi.org/10.1007/978-94-007-0449-7_1

Crawford, E. O., & Kirby, M. M. (2008). Fostering students' global awareness: technology applications in social studies teaching and learning. *Journal of Curriculum and Instruction, 2*(1), 56-73. https://doi.org/10.3776/joci.%25y.v2i1p56-73

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Heath and Company.

Durmaz, H., & Mutlu, S. (2014). The effects of an instructional intervention on 7th grade students' science process skills and science achievement. *Çukurova University Faculty of Education Journal*, *43*(2), 155. https://doi.org/10.14812/cufej.2014.018

Fisher, A. (2011). *Critical Thinking: An Introduction.* Cambridge University Press.

Gilbert, J. K., & Justi, R. (2016). *Modeling-based teaching in science education*. Springer.

Gürses, A., Cuya, Ş., Günes, K., & Doğar, Ç. (2014). Determination of the relation between undergraduate students' awareness levels regarding their scientific process skills and application potentials. *American Journal of Educational Research*, *2*(5), 250-256. https://doi.org/10.12691/education-2-5-3

Harrell, P. E. (2010). Teaching an integrated science curriculum: Linking teacher knowledge and teaching assignments. *Issues in teacher education*, *19*(1), 145-165.

Hidayati, T., Nugroho, S.E., & Sudarmin. (2013). Pengembangan tes diagnostik untuk mengidentifikasi keterampilan proses sains dengan tema energi pada pembelajaran ipa terpadu [Development of diagnostic tests to identify science process skills with the theme of energy in integrated science learning]. *Unnes Science Education Journal, 2*(2), 311-319. https://doi.org/10.15294/USEJ.V2I2.2041

Ismail, Z. H., & Jusoh, I. (2001). Relationship between science process skills and logical thinking abilities of Malaysian students. *Journal Of Science And Mathematics Education In Southeast Asia*, *24*(2), 67-77.

Joreskog, K. G., & Sobrorn, D. (1989) *LISREL 7 user's reference guide.* Scientifit Software International.

Karamustafaoğlu, S. (2011). Improving the science process skills ability of science student teachers using I diagrams. *International Journal of Physics & Chemistry Education*, *3*(1), 26-38. https://orcid.org/0000-0002-2852-7061

Karsli, F., & Şahin, Ç. (2009). Developing worksheet based on science process skills: Factors affecting solubility. *Asia-Pacific Forum on Science Learning and Teaching*, *10*(1), 1-12.

Kartowagiran, B., Suyanta, Hamdi, S., Jaedun, A., Ahman, Rusijono, & Laliyo, L. A. (2020). Development of web-based application for teacher candidate competence instruments: Preparing professional teachers in the IR 4.0 Era. *European Journal of Educational Research, 9*(4), 1749-1763. https://doi.org/10.12973/eu-jer.9.4.1749

Katchevich, D., Hofstein, A., & Mamlok-Naaman, R. (2013). Argumentation in the chemistry laboratory: Inquiry and confirmatory experiments. *Journal of Research in Science Education, 43*(1), 317-345. https://doi.org/10.1007/s11165-011-9267-9

Keil, C., Haney, J., & Zoffel, J. (2009). Improvements in student achievement and science process skills using environmental health science problem-based learning curricula. *The Electronic Journal for Research in Science & Mathematics Education, 13*(1), 1-18.

Ketterlin-Geller, L.R. (2005). Knowing what all students know: Procedures for developing universal design for assessment. *Journal of Technology, Learning, and Assessment, 4*(2), 1-22.

Kirsch, I., & Guthrie, J. (1980). Construct Validity of Functional Reading Tests. *Journal of Educational Measurement, 17*(2), 81-93.

Klein, J. D. (2014, April 3-7). *Design and development research: A rose by another name* [Paper presentation]. American Educational Research Association Conference (AERA) 2014, Philadelphia, PA, USA.

Koponen, I. T. (2007). Models and modelling in physics education: A critical re-analysis of philosophical underpinnings and suggestions for revisions. *Science & Education*, *16*(7-8), 751-773. doi.org/10.1007/s11191-006-9000-7

Kusiak, J., & Brown, D. (2007). *Creative Thinking.* Technique.

Leonor, J. P. (2015). Exploration of conceptual understanding and science process skills: A basis for differentiated science inquiry curriculum model. *International Journal of Information and Education Technology*, *5*(4), 255-259. https://doi.org/10.7763/IJIET.2015.V5.512

Lukum, A. (2015). Evaluasi program pembelajaran IPA SMP menggunakan model countenance stake [The evaluation of the SMP science learning program uses a stake countenance model]. *Journal of Educational Research and Evaluation/ Jurnal Penelitian dan Evaluasi Pendidikan*, *19*(1), 25-37. https://doi.org/10.21831/pep.v19i1.4552

Marzano, R. J. (2007). *The art of science teaching.* ASCD.

Munawaroh, H., Sudiyanto, & Riyadi. (2018). Teachers' perceptions of innovative learning model toward critical thinking ability. *International Journal of Educational Methodology*, *4*(3), 153-160. https://doi.org/10.12973/ijem.4.3.153

Özgelen, S. (2012). Students' science process skills within a cognitive domain framework. *Eurasia Journal of Mathematics, Science and Technology Education*, *8*(4), 283-292. https://doi.org/10.12973/eurasia.2012.846a

Ramadhan, S., Mardapi, D., Prasetyo, Z.K., & Utomo, H.B. (2019). The development of an instrument to measure the higher order thinking skill in physics. *European Journal of Educational Research, 8*(3), 743-751. https://doi.org/10.12973/eu-jer.8.3.743

Ramalingam, M., Kasilingam, G., & Chinnavan, E. (2014). Assessment of learning domains to improve student's learning in higher education. *Journal of Young Pharmacists, 6*(1), 27-33. https://doi.org/10.5530/jyp.2014.1.5

Resnick, L. B. (1987). *Education and learning to think*. National Academy Press.

Retnawati, H. (2016). *Validitas reliabilitas & karakteristik butir (panduan untuk peneliti, mahasiswa, dan psikometrian)* [Reliability validity & item characteristics (guide for researchers, students, and psychometrics)]. Parama Publishing.

Richey, R. C. (2005). Validating instruction design and development models. In J. M. Spector & D. A. Wiley (Eds.), *Innovations in instructional technology: Essays in honor of M. David Merrill* (pp. 171–185.) Lawrence Erlbaum Associates.

Richey, C. R. & Klein, J. D.(2007). *Design and development research*. Lawrence Erlbaum Associates.

Rotbain, Y., Marbach-Ad, G., & Stavy, R. (2006). Effect of bead and illustrations models on high school students' achievement in molecular genetics. *Journal of Research in Science Teaching, 43*(5), 500-529. https://doi.org/10.1002/tea.20144

Runco, M. A. (1993). *Creativity as an educational objective for disadvantaged students*. National Research Center on the Gifted and Talented, The University of Connecticut. https://eric.ed.gov/?id=ED363074

South African Qualifications Authority. (2014). *National policy and criteria for designing and implementing assessment for NQF qualifications and part-qualifications and professional designations in South Africa*. https://cutt.ly/bjsrwox

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research*, *8*(2), 23-74.

Shavelson, R., & Stanton, G. (1975). Construct validation: Methodology and application to three measures of cognitive structure. *Journal of Educational Measurement*, *12*(2), 67-85.

Shirban, S, A. (2018). Enhancing critical thinking in Malaysian primary school students through PLS method. *International Journal of Educational Methodology, 4*(4), 243-257. https://doi.org/10.12973/ijem.4.4.243

Spector, J. M., Merrill, M. D., Ellen, J., & Bishop, M. J. (Eds.). (2014). *Handbook of research on educational and technology*. Springer.

Subali, B. (2009). *Pengukuran keterampilan proses sains pola divergen dalam mata pelajaran biologi SMA di Provinsi DIY dan Jawa Tengah* [Measurement of divergent pattern science process skills in high school biology subjects in Yogyakarta and Central Java provinces] [Unpublished doctoral dissertation]. Yogyakarta State University.

Temiz, B. K., Taşar, M. F., & Tan, M. (2006). Development and validation of a multiple format test of science process skills. *International Education Journal*, *7*(7), 1007-1027.

Tracey, M. W., & Richey, R. C. (2007). ID model construction and validation: A multiple intelligences case. *Educational Technology Research and Development*, *55*(4), 369-390. https://doi.org/10.1007/s11423-006-9015-4